# **MyOnlineShop Internship Program**

**Team Ace** **Week 7: Advanced Cohort Analysis – Weekly Retention & Churn Trends**

**Timeline**: Two Weeks (Week 7-8)  
**Submission:** Wednesday, 3rd September 2025 (4pm)

## **Getting Started: Why This Week Matters**

This week, we’re introducing you to a concept that plays a big role in how businesses, especially digital platforms, understand their customers: **cohort analysis**. Think of it as a smarter way to track and compare how different groups of users behave over time, rather than looking at everyone in one big lump.

You’ll often hear about big, shiny numbers like total users or monthly signups. While they might look impressive, those surface-level stats rarely tell the full story. They don’t show how engaged users are or whether they actually stick around. That’s where cohort analysis comes in, it breaks users into groups based on a common trait (like sign-up month) and lets us track how long each group stays active, how they engage, and where they fall off.

Understanding this behavior is key to making better product decisions, improving customer experiences, and increasing long-term value. You’ll also dive into two powerful companion metrics, **retention** and **churn**, which help us measure who’s sticking around and who’s leaving, and why.

## **The Concepts, Simplified**

* **Cohort Analysis** is a way to group users by shared traits, like the month they signed up, and observe how their behavior evolves. This gives us a time-based view of user engagement, highlighting trends that aren’t obvious in total numbers.
* Imagine two customers: one who signed up during a major Black Friday promotion and another who joined during a quiet off-season month. Even if both signed up, their likelihood to stick around might be very different. With cohort analysis, we can track those differences.
* **Retention Rate** tells us how many users are still active after a certain time period. It’s a signal of long-term engagement, the kind of thing that keeps a business sustainable.
* **Churn Rate**, on the other hand, tells us how many users dropped off, those who stopped using the product or canceled their subscription. It’s often more alarming, but just as valuable. As Bill Gates once said, *“Your most unhappy customers are your greatest source of learning.”*
* The better you understand these two sides of user behavior, the more effectively you can help reduce churn and build loyalty.

## **Learning Resources to Explore First**

Start by familiarizing yourself with the fundamentals of cohort analysis. These materials will walk you through the theory and help you see how different teams, from marketers to product managers, use these insights to improve user experience.

1. [**What is Cohort Analysis?**](https://pg-p.ctme.caltech.edu/blog/data-analytics/what-is-cohort-analysis-types-benefits-steps)  [Read the blog post by Tomi Mester](https://www.retention.blog/p/why-should-i-care-about-cohort-analysis)
2. [**Retention vs. Churn Explained**](https://posthog.com/product-engineers/churn-rate-vs-retention-rate) [Understanding Retention vs Churn Metrics](https://data-science-churn.medium.com/how-to-calculate-churn-with-sql-fighting-churn-with-data-f4e5fb13a4df)
3. [**Customer Retention and Cohort Analysis**](https://www.youtube.com/watch?v=OwCATJh4lNg)

## **Project Overview**

This week, we’re taking a step further by focusing on weekly cohort retention, an essential skill for analysts in fast-moving subscription-based businesses. Your Product Manager (PM) has asked for a deeper look into churn and retention trends. Monthly retention, while useful, may delay key insights. Instead, you'll calculate weekly retention rates across user cohorts and track how long users remain active over the first 6 weeks of their subscription. Assume that you are doing this analysis on 2021-02-07

This project will require working directly with a database, writing custom SQL, and visualizing the resulting data using tools such as Google Sheets, Power BI, or Tableau. By the end of this task, you should be able to deliver a clear, actionable picture of user retention trends across time.

## **Objective**

To calculate and visualize weekly retention rates for user cohorts, identify trends in user engagement and churn, and communicate insights in a visual and analytical format that supports data-driven decision-making.

## **Steps to Follow**

### **1. Understand the Data:** begin by familiarizing yourself with the structure of the subscriptions table. Focus on:

* user\_pseudo\_id
* subscription\_start\_date
* subscription\_end\_date

Use an exploratory query to check for missing data, duplicate records, and unexpected values. Make note of anything that could affect your retention calculations.

### **2. Start Small: One Cohort:** choose a weekly cohort, for example, users who began subscriptions from **2021-01-04 to 2021-01-11**. Use SQL to:

* Truncate the subscription\_start\_date to the start of the week.
* Track whether these users were still active in Week 1 through Week 6 (based on their subscription end date).
* Calculate the **weekly retention rate** by comparing how many users from the original cohort were still active in each subsequent week.

This will serve as your foundation before scaling up the query.

### **3. Extend to Multiple Cohorts:** modify your SQL to calculate retention rates across all weekly cohorts (not just one). Your output should show:

* Each cohort start week (as cohort\_week)
* Retention rates for Week 0 (initial sign-up) through Week 6

Use conditional aggregation to count how many users in each cohort were active in each week, then divide by the original cohort size to calculate retention percentages.

### **4. Export the Data:** once your SQL query returns the desired results:

* Export your dataset as a CSV, or
* Use BigQuery’s **"Open In"** feature to send your results directly to Google Sheets or your chosen BI tool.

[**Exported Data**](https://docs.google.com/spreadsheets/d/1f7apxoNSSavG-7qwqUUxlk05tG5deXiVqazgiGe1i8A/edit?usp=sharing)

### **5. Visualize Retention Trends:** use a heatmap to visualize weekly retention by cohort. Each row should represent a cohort (e.g., “Week of 2021-01-04”), and each column should represent a week (Week 0 to Week 6). Use color gradients (green to red) to indicate high to low retention. You may also use line or bar charts to illustrate trends within individual cohorts over time.

### **6. Interpret & Summarize Your Insights:** once your visualizations are complete:

* Identify where drop-offs are most severe
* Highlight strong-performing cohorts (e.g., where retention is unusually high)
* Offer recommendations: Could the drop-off be related to pricing? Lack of onboarding? Marketing timing?

You should also reflect on how cohort size, acquisition timing, or user intent may explain differences in retention.

## **Final Deliverables**

1. **Cohort Retention Dataset** Exported data from your SQL query showing retention rates for each cohort (Week 0 to Week 6).
2. **Retention Visualization (Heatmap or Chart)** A clear and color-coded visualization of weekly retention by cohort, created in Google Sheets, Power BI, Tableau, or another visualization tool.
3. **Insights & Summary Document (1–2 pages)** A written explanation of your:
   * Methodology
   * Observations and trends
   * Key insights and recommendations
   * Sample SQL snippets (optional but encouraged)

## **A Few Reminders**

This is a real-world type of project that many product and growth analysts tackle daily. Your ability to track retention weekly instead of monthly means your insights can influence decisions faster and help prevent customer churn earlier.

Also, keep in mind the balance between retention and churn. A dip in one typically implies a spike in the other, so always view your results through that dual lens.

#### **Submission Details:**

Each group must submit their report via email to [**gabrileyashim405@gmail.com**](mailto:gabrileyashim405@gmail.com) .

**File Naming Format:**<Team name>-<Week number>: <Group name>

**Example:**Team Ace-Week 5: Group Ace

**Important:**

* Ensure the document is set to **“Anyone with the link can view”** before submitting.
* Ensure to list the name of every Group member that participated, and copy them on the email.
* The email **subject line** should be: **"Team Ace - Week 5 Submission: [Full Name]"**.